

1646

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/276,455 DATE: 05/11/2000
 TIME: 15:43:20

Input Set : A:\Fershtl.app
 Output.Set: N:\CRF3\05112000\I276455.raw

3 <110> APPLICANT: FERSHT, Alan R.
 4 ZAHN, Ralph
 5 ALTAMIRANO, Myriam M.
 7 <120> TITLE OF INVENTION: CHAPERONE FRAGMENTS
 9 <130> FILE REFERENCE: 674508-2001
 11 <140> CURRENT APPLICATION NUMBER: 09/276,455
 12 <141> CURRENT FILING DATE: 1999-03-25
 14 <150> PRIOR APPLICATION NUMBER: 97/02652
 15 <151> PRIOR FILING DATE: 1997-09-26
 17 <150> PRIOR APPLICATION NUMBER: 96/020980
 18 <151> PRIOR FILING DATE: 1996-12-03
 20 <150> PRIOR APPLICATION NUMBER: 9620243.7
 21 <151> PRIOR FILING DATE: 1996-09-26
 23 <160> NUMBER OF SEQ ID NOS: 14
 25 <170> SOFTWARE: PatentIn Ver. 2.1
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 27
 29 <212> TYPE: DNA
 30 <213> ORGANISM: Escherichia coli
 32 <400> SEQUENCE: 1
 33 cggatccgaa ggtatgcagt tcgaccg 27
 36 <210> SEQ ID NO: 2
 37 <211> LENGTH: 30
 38 <212> TYPE: DNA
 39 <213> ORGANISM: Escherichia coli
 41 <400> SEQUENCE: 2
 42 cgaattctta aacgccgcct gccagtttcg 30
 45 <210> SEQ ID NO: 3
 46 <211> LENGTH: 31
 47 <212> TYPE: DNA
 48 <213> ORGANISM: Escherichia coli
 50 <400> SEQUENCE: 3
 51 cgaattctta acggccctgg attgcagctt c 31
 54 <210> SEQ ID NO: 4
 55 <211> LENGTH: 35
 56 <212> TYPE: DNA
 57 <213> ORGANISM: Escherichia coli
 59 <400> SEQUENCE: 4
 60 cgaattctta acccacgccca tcgatgatag tgggtg 35
 63 <210> SEQ ID NO: 5
 64 <211> LENGTH: 42
 65 <212> TYPE: DNA
 66 <213> ORGANISM: Escherichia coli
 68 <400> SEQUENCE: 5
 69 cgaattctta gtctttgttg atcacaacac gtttagcctg ac 42
 72 <210> SEQ ID NO: 6
 73 <211> LENGTH: 36

ENTERED

RAW SEQUENCE LISTING DATE: 05/11/2000
 PATENT APPLICATION: US/09/276,455 TIME: 15:43:20
 Input Set : A:\Fersht1.app
 Output Set: N:\CRF3\05112000\I276455.raw

```

74 <212> TYPE: DNA
75 <213> ORGANISM: Escherichia coli
77 <400> SEQUENCE: 6
78 cgaattctta acgttttagcc tgaccacaggt ctccca           36
81 <210> SEQ ID NO: 7
82 <211> LENGTH: 34
83 <212> TYPE: DNA
84 <213> ORGANISM: Escherichia coli
86 <400> SEQUENCE: 7
87 cgaattctta accgccagtc agggttgcga tatc               34
90 <210> SEQ ID NO: 8
91 <211> LENGTH: 36
92 <212> TYPE: PRT
93 <213> ORGANISM: Escherichia coli
95 <400> SEQUENCE: 8
96 Met Arg Gly Ser His His His His His His Gly Met Ala Ser Met Thr
97   1           5           10          15
99 Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Leu Val
100   20          25          30
102 Pro Arg Gly Ser
103   35
106 <210> SEQ ID NO: 9
107 <211> LENGTH: 186
108 <212> TYPE: PRT
109 <213> ORGANISM: Escherichia coli
111 <400> SEQUENCE: 9
112 Glu Gly Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Asn
113   1           5           10          15
115 Lys Pro Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile Leu Leu
116   20          25          30
118 Ala Asp Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val Leu Glu
119   35          40          45
121 Ala Val Ala Lys Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Val
122   50          55          60
124 Glu Gly Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg Gly Ile
125  65          70          75          80
127 Val Lys Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys
128   85          90          95
130 Ala Met Leu Gln Asp Ile Ala Thr Leu Thr Gly Gly Thr Val Ile Ser
131  100         105         110
133 Glu Glu Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp Leu Gly
134  115         120         125
136 Gln Ala Lys Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile Ile Asp
137  130         135         140
139 Gly Val Gly Glu Glu Ala Ala Ile Gln Gly Arg Val Ala Gln Ile Arg
140 145         150         155         160
142 Gln Gln Ile Glu Glu Ala Thr Ser Asp Tyr Asp Arg Glu Lys Leu Gln
143   165         170         175
145 Glu Arg Val Ala Lys Leu Ala Gly Gly Val

```

RAW SEQUENCE LISTING

DATE: 05/11/2000

PATENT APPLICATION: US/09/276,455

TIME: 15:43:20

Input Set : A:\Fershtl.app

Output Set: N:\CRF3\05112000\I276455.raw

```

146          180          185
149 <210> SEQ ID NO: 10
150 <211> LENGTH: 548
151 <212> TYPE: PRT
152 <213> ORGANISM: Escherichia coli
154 <400> SEQUENCE: 10
155 Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala Arg Val Lys Met
156 1 5 10 15
158 Leu Arg Gly Val Asn Val Leu Ala Asp Ala Val Lys Val Thr Leu Gly
159 20 25 30
161 Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe Gly Ala Pro Thr
162 35 40 45
164 Ile Thr Lys Asp Gly Val Ser Val Ala Arg Glu Ile Glu Leu Glu Asp
165 50 55 60
167 Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Ser Lys
168 65 70 75 80
170 Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala
171 85 90 95
173 Gln Ala Ile Ile Thr Glu Gly Leu Lys Ala Val Ala Ala Gly Met Asn
174 100 105 110
176 Pro Met Asp Leu Lys Arg Gly Ile Asp Lys Ala Val Thr Ala Ala Val
177 115 120 125
179 Glu Glu Leu Lys Ala Leu Ser Val Pro Cys Ser Asp Ser Lys Ala Ile
180 130 135 140
182 Ala Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Glu Thr Val Gly Lys
183 145 150 155 160
185 Leu Ile Ala Glu Ala Met Asp Lys Val Gly Lys Glu Gly Val Ile Thr
186 165 170 175
188 Val Glu Asp Gly Thr Gly Leu Gln Asp Glu Leu Asp Val Val Glu Gly
189 180 185 190
191 Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Asn Lys Pro
192 195 200 205
194 Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile Leu Leu Ala Asp
195 210 215 220
197 Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val Leu Glu Ala Val
198 225 230 235 240
200 Ala Lys Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly
201 245 250 255
203 Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg Gly Ile Val Lys
204 260 265 270
206 Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met
207 275 280 285
209 Leu Gln Asp Ile Ala Thr Leu Thr Gly Gly Thr Val Ile Ser Glu Glu
210 290 295 300
212 Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp Leu Gly Gln Ala
213 305 310 315 320
215 Lys Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile Ile Asp Gly Val
216 325 330 335
218 Gly Glu Glu Ala Ala Ile Gln Gly Arg Val Ala Gln Ile Arg Gln Gln

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/276,455

DATE: 05/11/2000
TIME: 15:43:20

Input Set : A:\Pershtl.app
Output Set: N:\CRF3\05112000\I276455.raw

```

219           340           345           350
221 Ile Glu Glu Ala Thr Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg
222           355           360           365
224 Val Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala
225           370           375           380
227 Thr Glu Val Glu Met Lys Glu Lys Lys Ala Arg Val Glu Asp Ala Leu
228 385           390           395           400
230 His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val Ala Gly Gly Gly
231           405           410           415
233 Val Ala Leu Ile Arg Val Ala Ser Lys Leu Ala Asp Leu Arg Gly Gln
234           420           425           430
236 Asn Glu Asp Gln Asn Val Gly Ile Lys Val Ala Leu Arg Ala Met Glu
237           435           440           445
239 Ala Pro Leu Arg Gln Ile Val Leu Asn Cys Gly Glu Glu Pro Ser Val
240           450           455           460
242 Val Ala Asn Thr Val Lys Gly Gly Asp Gly Asn Tyr Gly Tyr Asn Ala
243 465           470           475           480
245 Ala Thr Glu Glu Tyr Gly Asn Met Ile Asp Met Gly Ile Leu Asp Pro
246           485           490           495
248 Thr Lys Val Thr Arg Ser Ala Leu Gln Tyr Ala Ala Ser Val Ala Gly
249           500           505           510
251 Leu Met Ile Thr Thr Glu Cys Met Val Thr Asp Leu Pro Lys Asn Asp
252           515           520           525
254 Ala Ala Asp Leu Gly Ala Ala Gly Gly Met Gly Gly Met Gly Gly Met
255           530           535           540
257 Gly Gly Met Met
258 545
261 <210> SEQ ID NO: 11
262 <211> LENGTH: 160
263 <212> TYPE: PRT
264 <213> ORGANISM: Escherichia coli
266 <400> SEQUENCE: 11
267 Val Glu Gly Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Ile
268 1           5           10           15
270 Asn Lys Pro Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile Leu
271           20           25           30
273 Leu Ala Asp Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val Leu
274           35           40           45
276 Glu Ala Val Ala Lys Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp
277           50           55           60
279 Val Glu Gly Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg Gly
280 65           70           75           80
282 Ile Val Lys Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg
283           85           90           95
285 Lys Ala Met Leu Gln Asp Ile Ala Thr Leu Thr Gly Gly Thr Val Ile
286           100          105          110
288 Ser Glu Glu Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp Leu
289           115          120          125
291 Gly Gln Ala Lys Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile Ile

```

RAW SEQUENCE LISTING DATE: 05/11/2000
 PATENT APPLICATION: US/09/276,455 TIME: 15:43:20

Input Set : A:\Fershtl.app
 Output Set: N:\CRF3\05112000\I276455.raw

```

292      130      135      140
294 Asp Gly Val Gly Glu Glu Ala Ala Ile Gln Gly Arg Val Ala Gln Ile
295 145      150      155      160
301 <210> SEQ ID NO: 12
302 <211> LENGTH: 42
303 <212> TYPE: PRT
304 <213> ORGANISM: Escherichia coli
306 <220> FEATURE:
307 <221> NAME/KEY: VARIANT
308 <222> LOCATION: (1)
309 <223> OTHER INFORMATION: Amino acid 1 is Ile, Met, Leu, Val, Ser, Phe or
310 Ala
312 <220> FEATURE:
313 <221> NAME/KEY: VARIANT
314 <222> LOCATION: (5)
315 <223> OTHER INFORMATION: Amino acid 5 is Ile, Leu, Val, Pro or Ala
317 <220> FEATURE:
318 <221> NAME/KEY: VARIANT
319 <222> LOCATION: (8)
320 <223> OTHER INFORMATION: Amino acid 8 is Leu, Glu, Val, His or Ile
322 <220> FEATURE:
323 <221> NAME/KEY: VARIANT
324 <222> LOCATION: (9)
325 <223> OTHER INFORMATION: Amino acid 9 is Glu, Ala, Arg, Leu, Gln or Asn
327 <220> FEATURE:
328 <221> NAME/KEY: VARIANT
329 <222> LOCATION: (12)
330 <223> OTHER INFORMATION: Amino acid 12 is Ala, Val, Ile, Met, Leu, Asn,
331 Ser, Arg, Thr, Gln or Lys
333 <220> FEATURE:
334 <221> NAME/KEY: VARIANT
335 <222> LOCATION: (28)
336 <223> OTHER INFORMATION: Amino acid 28 is Glu, Asp or Gly
338 <220> FEATURE:
339 <221> NAME/KEY: VARIANT
340 <222> LOCATION: (31)
341 <223> OTHER INFORMATION: Amino acid 31 is Ala, Pro, Ser, Thr, Gly, or Leu
343 <220> FEATURE:
344 <221> NAME/KEY: VARIANT
345 <222> LOCATION: (32)
346 <223> OTHER INFORMATION: Amino acid 32 is Thr, Ala, Asn, Ser, or Val
348 <220> FEATURE:
349 <221> NAME/KEY: VARIANT
350 <222> LOCATION: (34)
351 <223> OTHER INFORMATION: Amino acid 34 is Val, Leu, Ile, or Ala
353 <220> FEATURE:
354 <221> NAME/KEY: VARIANT
355 <222> LOCATION: (35)
356 <223> OTHER INFORMATION: Amino acid 35 is Val, Leu, Ile, Phe or His

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

DATE: 05/11/2000

PATENT APPLICATION: US/09/276,455

TIME: 15:43:21

Input Set : A:\Pershtl.app

Output Set: N:\CRF3\05112000\I276455.raw

L:420 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:423 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:426 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:441 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:444 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:447 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13